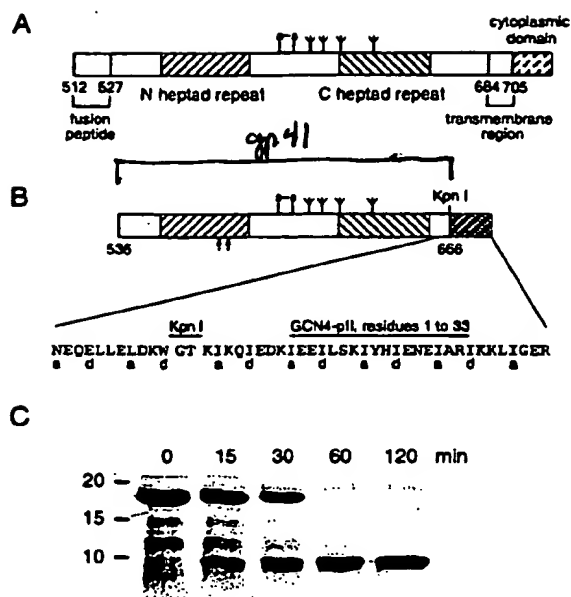


09877505-060904



gpn41
536 — 666 GNM

NATURE 371:80
1994

HARBOY P.A.

Figure 2

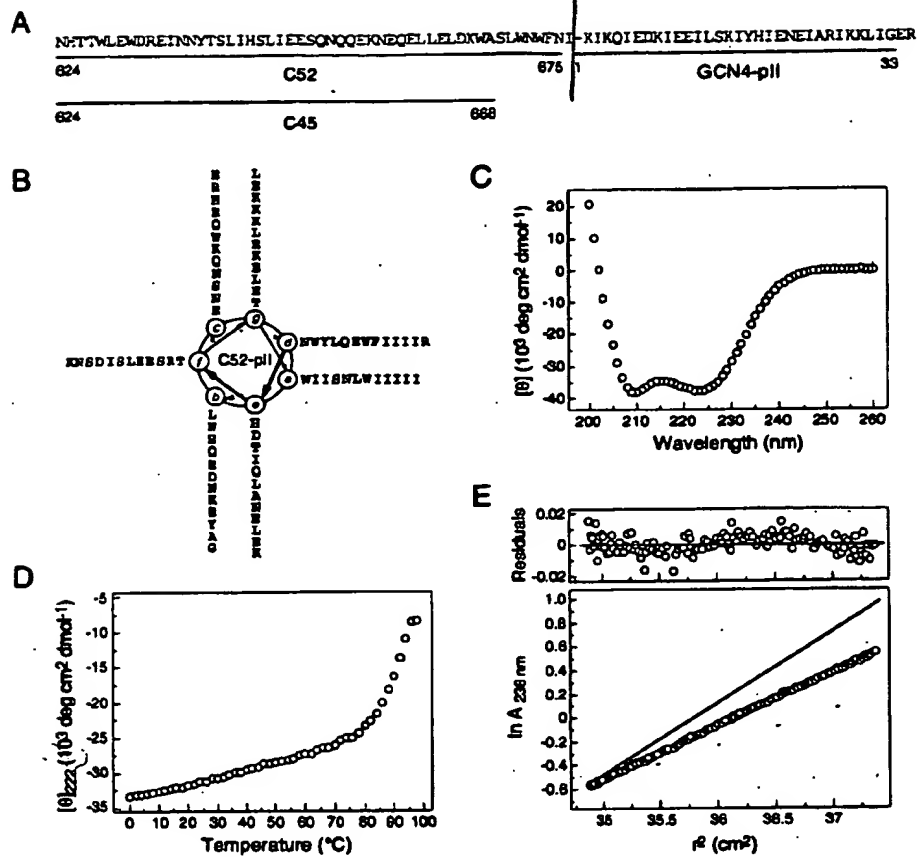


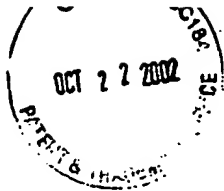
Figure 7

00877505-050001

| | | | | | | |
|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| 669 | 670 | 671 | 672 | 673 | 674 | 675 |
| L ²⁰⁵ | W ²¹⁰ | N ¹⁶³ | W ²¹² | F ²⁰⁴ | D ¹²⁹ | I ²¹² |
| I ⁶ | • 3 | S ³⁸ | • 1 | L ⁷ | S ⁴⁶ | M ¹ |
| V ¹ | | T ¹⁰ | | Y ¹ | N ²³ | |
| S ¹ | | D ¹ | | S ¹ | G ³ | |
| | | | | | T ³ | |
| | | | | | E ² | |
| | | | | | • 1 | |

gplao

Figure 8



SEQUENCE LISTING

COPY

<110> Lu, Min
Ji, Hong

<120> Antigen for Developing Neutralizing Antibodies to Human Immunodeficiency
Virus

<130> 1676.002US1

<140> 09/877,606

<141> 2001-06-08

<150> US 60/210,322

<151> 2000-06-08

<160> 9

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 46

<212> PRT

<213> Artificial Sequence

<220>

<223> The amino acid sequence of the COOH-terminal extension of gp41.

<400> 1

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Glu | Gln | Glu | Leu | Leu | Glu | Leu | Asp | Lys | Trp | Gly | Thr | Lys | Ile | Lys |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Gln | Ile | Glu | Asp | Lys | Ile | Glu | Glu | Ile | Leu | Ser | Lys | Ile | Tyr | His | Ile |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Glu | Asn | Glu | Ile | Ala | Arg | Ile | Lys | Lys | Leu | Ile | Gly | Glu | Arg | | |
| | | 35 | | | | | 40 | | | | | 45 | | | |

<210> 2

<211> 78

<212> PRT

<213> Artificial Sequence

<220>

<223> Linear amino acid sequence of a helical wheel projection of C45-pII.

<400> 2

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | His | Thr | Thr | Trp | Leu | Glu | Trp | Asp | Arg | Glu | Ile | Asn | Asn | Tyr | Thr |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |
| Ser | Leu | Ile | His | Ser | Leu | Ile | Glu | Glu | Ser | Gln | Asn | Gln | Gln | Glu | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Asn | Glu | Gln | Glu | Leu | Leu | Glu | Leu | Asp | Lys | Trp | Ala | Ser | Lys | Ile | Lys |
| | | 35 | | | | | 40 | | | | 45 | | | | |
| Gln | Ile | Glu | Asp | Lys | Ile | Glu | Glu | Ile | Leu | Ser | Lys | Ile | Tyr | His | Ile |
| | | 50 | | | | 55 | | | | 60 | | | | | |
| Glu | Asn | Glu | Ile | Ala | Arg | Ile | Lys | Lys | Leu | Ile | Gly | Glu | Arg | | |
| 65 | | | | | 70 | | | | | 75 | | | | | |

COPY

<210> 3
<211> 52
<212> PRT
<213> Artificial Sequence

<220>
<223> Amino acid sequence of C52.

<400> 3
Asn His Thr Thr Trp Leu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr
1 5 10 15
Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys
20 25 30
Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn
35 40 45
Trp Phe Asn Ile
50

<210> 4
<211> 85
<212> PRT
<213> Artificial Sequence

<220>
<223> The chimera of the C52 peptide and an isoleucine-zipper trimer (GCN4-pII).

<400> 4
Asn His Thr Thr Trp Leu Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr
1 5 10 15
Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys
20 25 30
Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn
35 40 45
Trp Phe Asn Ile Lys Ile Lys Gln Ile Glu Asp Lys Ile Glu Glu Ile
50 55 60
Leu Ser Lys Ile Tyr His Ile Glu Asn Glu Ile Ala Arg Ile Lys Lys
65 70 75 80
Leu Ile Gly Glu Arg
85

<210> 5
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Most common sequence of amino acids at positions 669-675 of gp160 (M group HIV-1 strain).

<400> 5
Leu Trp Asn Trp Phe Asp Ile
1 5

<210> 6
<211> 6
<212> PRT
<213> Artificial Sequence

<220>

<223> Amino acid sequence immediately NH2-terminal to SEQ ID NO:5.

<400> 6

Glu Leu Asp Lys Trp Ala

1

5

<210> 7

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Amino acids 669-675 of gp41.

<400> 7

Leu Trp Asn Trp Phe Asn Ile

1

5

<210> 8

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Amino acids at positions 669-675 of gp160.

<221> SITE

<222> 1

<223> Xaa = Leu, Ile, Val, or Ser.

<221> SITE

<222> 2

<223> Xaa = tryptophan or no amino acid present due to the presence of a stop codon at this position in the reading frame.

<221> SITE

<222> 3

<223> Xaa = Asn, Ser, Thr, or Asp or no amino acid present due to the presence of a stop codon at an earlier position in the reading frame.

<221> SITE

<222> 4

<223> Xaa = Trp or no amino acid present due to the presence of a stop codon at this position or at an earlier position in the reading frame.

<221> SITE

<222> 5

<223> Xaa = Phe, Leu, Tyr, or Ser or no amino acid present due to the presence of a stop codon at an earlier position in the reading frame.

<221> SITE

<222> (6)...(6)

<223> Xaa = Asp, Ser, Asn, Gly, Thr, Glu, or no amino acid present due to the presence of a stop codon at this position or an earlier position in the reading frame.

<221> SITE

<222> (7)...(7)

<223> Xaa = Ile or Met or no amino acid present due to the presence of a stop codon at an earlier position in the reading frame.

<400> 8

Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5

<210> 9

<211> 52

<212> PRT

<213> Artificial Sequence

<220>

<223> Mutant (N656L) amino acid sequence of C52.

<400> 9

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | His | Thr | Thr | Trp | Leu | Glu | Trp | Asp | Arg | Glu | Ile | Asn | Asn | Tyr | Thr |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |
| Ser | Leu | Ile | His | Ser | Leu | Ile | Glu | Glu | Ser | Gln | Asn | Gln | Gln | Glu | Lys |
| | | | 20 | | | | 25 | | | | | | 30 | | |
| Leu | Glu | Gln | Glu | Leu | Leu | Glu | Leu | Asp | Lys | Trp | Ala | Ser | Leu | Trp | Asn |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Trp | Phe | Asn | Ile | | | | | | | | | | | | |
| | | | 50 | | | | | | | | | | | | |

Issue Classification

(Assistant Examiner) (Date)
C. E. Phelan 06/28/04
(Legal Instruments Examiner) (Date)

☐ Claims renumbered in the same order as presented by applicant



Helical Interactions in the HIV-1 gp41 Core Reveal Structural Basis for the Inhibitory Activity of gp41 Peptides^{†,‡}

Wei Shu,[§] Jie Liu,[§] Hong Ji,[§] Lin Radigen,^{||} Shibo Jiang,^{||} and Min Lu^{*,§}

Department of Biochemistry, Weill Medical College of Cornell University, 1300 York Avenue, New York, New York 10021, and Lindsley F. Kimball Research Institute, New York Blood Center, New York, New York 10021

Received September 20, 1999; Revised Manuscript Received November 30, 1999

ABSTRACT: The HIV-1 gp41 envelope protein mediates membrane fusion that leads to virus entry into the cell. The core structure of fusion-active gp41 is a six-helix bundle in which an N-terminal three-stranded coiled coil is surrounded by a sheath of antiparallel C-terminal helices. A conserved glutamine (Gln 652) buried in this helical interface replaced by leucine increases HIV-1 infectivity. To define the basis for this enhanced membrane fusion activity, we investigate the role of the Gln 652 to Leu substitution on the conformation, stability, and biological activity of the N34(L6)C28 model of the gp41 ectodomain core. The 2.0 Å resolution crystal structure of the mutant molecule shows that the Leu 652 side chains make prominent contacts with hydrophobic grooves on the surface of the central coiled coil. The Gln 652 to Leu mutation leads to a marginal stabilization of the six-helix bundle by −0.8 kcal/mol, evaluated from thermal unfolding experiments. Strikingly, the mutant N34(L6)C28 peptide is a potent inhibitor of HIV-1 infection, with 10-fold greater activity than the wild-type molecule. This inhibitory potency can be traced to the corresponding C-terminal mutant peptide that likely has greater potential to interact with the coiled-coil trimer. These results provide strong evidence that conserved interhelical packing interactions in the gp41 core are important determinants of HIV-1 entry and its inhibition. These interactions also offer a test-bed for the development of more potent analogues of gp41 peptide inhibitors.

The human immunodeficiency virus type 1 (HIV-1)¹ enters cells by the envelope protein-mediated membrane fusion process that allows invasion of the infectious viral genome. The envelope glycoprotein of HIV-1 consists of a complex of the surface subunit gp120 and the transmembrane subunit gp41, which are generated through proteolytic cleavage of the polypeptide precursor gp160 (reviewed in refs 1–3). The envelope glycoprotein is displayed on the lipid bilayer of the virion as well as on the plasma membrane of infected cells. gp120 is responsible for binding to the CD4 glycoprotein and a chemokine receptor on the T cell surface (4, 5). Subsequently, gp41 undergoes conformational changes that facilitate membrane fusion between virus and target cell during infection, and between infected cells and uninfected cells during multinucleated giant cell (syncytium) formation (3). This receptor binding-promoted fusion activation is thought to involve conformational changes in the envelope

protein complex from a native (nonfusogenic) to a fusion-active (fusogenic) state (for a recent review, see ref 6). The control of this fusion activation process involving multiple protein interactions is central to the complex biology of HIV-1 entry and its inhibition.

The gp41 molecule contains several characteristic features within its ectodomain. Toward its amino terminus is the hydrophobic, glycine-rich sequence referred to as the fusion peptide, which is thought to insert into the bilayer of the target membrane at an early step of the fusion process (7–11). This fusion peptide region is followed by two regions that exhibit a 4–3 repeat of hydrophobic amino acids with the potential to form α -helical coiled-coil structures (12–14). The N (amino)-terminal heptad repeat is located adjacent to the fusion peptide, while the C (carboxyl)-terminal repeat precedes the transmembrane segment. Between these heptad-repeat regions is a disulfide-bonded loop region. Limited proteolysis studies identified a stable core of the gp41 ectodomain, in which the two heptad-repeat regions form a trimer of antiparallel dimers (15, 16). The X-ray crystal structure of this core reveals that three N-terminal helices form a central, three-stranded coiled coil, whereas three C-terminal helices pack in an antiparallel orientation into three hydrophobic grooves on the surface of this coiled-coil trimer (17–19). We have previously shown that a monoclonal antibody specifically recognizing this six-helix bundle binds to the surface of HIV-1 infected cells only after interaction of the envelope protein complex with soluble CD4 (20). This finding provides direct evidence for the notion that the gp41 ectodomain core represents the fusion-active conformation of the HIV-1 envelope protein

[†] This work was supported by National Institutes of Health Grants AI42693 to S.J. and AI42382 to M.L. and by the New York City Council Speaker's Fund for Biomedical Research to M.L.

[‡] The refined coordinates have been deposited in the Protein Data Bank with accession number 1DLB.

* To whom correspondence should be addressed. Phone: (212) 746-6562. Fax: (212) 746-8875. E-mail: mlu@mail.med.cornell.edu.

[§] Weill Medical College of Cornell University.

^{||} Lindsley F. Kimball Research Institute.

¹ Abbreviations: HIV-1, human immunodeficiency virus type 1; $[\theta]_{222}$, molar ellipticity at 222 nm; CD, circular dichroism; HPLC, high-performance liquid chromatography; LB, Luria-Bertani; T_m , midpoint of thermal denaturation; GdmCl, guanidinium hydrochloride; IPTG, isopropylthio- β -galactoside; PBS, neutral pH phosphate-buffered saline; PDB, Protein Data Bank; rms, root-mean-square; BCECF-AM, 2',7'-bis-(2-carboxyethyl)-5-(and-6)-carboxyfluorescein-acetoxymethyl ester; CPE, HIV-1_{IIIB}-mediated cytopathic effect.

2000/1
TRACH
CL-1000
FUSION PRO.

(15, 17–19). In addition, the six-helix bundle appears to play a direct role in membrane fusion because the folding and stability of the gp41 core in vitro correlate well with severity of the in vivo phenotypes observed in cells expressing engineered HIV-1 envelope protein variants (21, 22).

The current model for gp41-mediated membrane fusion is that a conformational change and subsequent antiparallel association of α -helices in the gp41 ectodomain place the viral and cellular membranes into close proximity for fusion (18, 23, 24). According to this model, packing interactions between the central coiled-coil trimer and the C-terminal helix are critical for membrane fusion. In the crystal structure of the gp41 core, Gln 652 and Asn 656 at *a* and *d* positions of the C-terminal helix pack against the N-terminal coiled coil (17–19). When Gln 652 was replaced by leucine, the mutant virus is more infectious than the wild-type virus (25). In contrast, a leucine substitution for Asn 656 abolishes infectivity and membrane fusion (25). Thus, the buried face of the C-terminal helix and its packing interactions with the N-terminal coiled coil appear to control the key process of HIV-1 entry. These conserved packing interactions also play a major role in the ability of synthetic peptides corresponding to the C-terminal helix to inhibit HIV-1 infection (26, 27). The available evidence suggests that these peptide inhibitors act in a dominant-negative manner by binding to the N-terminal coiled coil of gp41 (15, 17–19, 24, 26–32). In addition, recent studies indicate that the structural and mechanistic basis of helical interactions is conserved in many viral membrane fusion proteins (e.g., refs 33 and 34).

Here we use the recombinant N34(L6)C28 model (21) to characterize the effects of the Gln 652 to Leu mutation on the structure and biological activity of the gp41 ectodomain core. Circular dichroism and sedimentation equilibrium measurements indicate that the Q652L mutant forms a trimeric helical structure. The stability of the Q652L trimer increases by 0.8 kcal/mol, relative to the wild-type complex. The X-ray crystal structure of Q652L at 2.0 Å resolution reveals that the mutant sequence is accommodated in the six-helix bundle by forming a packing surface with different sets of atoms. Therefore, the Leu substitution for Gln 652 strengthens the interhelix interaction by providing additional hydrophobic packing forces and/or removing the burial of polar atoms. Moreover, the Q652L mutant peptide is a potent inhibitor of HIV-1 infection, with 10-fold greater activity than the wild-type peptide. This inhibitory potency correlates with the coiled-coil-binding potential of the corresponding C-terminal mutant peptide. Thus, the packing interactions of the six-helix bundle are likely to underlie the mechanism of membrane fusion. They also provide a good test-bed for developing more potent versions of C peptides.

MATERIALS AND METHODS

Mutagenesis and Peptide Production. Mutations were introduced into the N34(L6)C28 model by single-stranded mutagenesis (35) and verified by DNA sequencing. All recombinant peptides were expressed in *Escherichia coli* BL21(DE3) pLysS using the T7 expression system (36). Cells, freshly transformed with an appropriate plasmid, were grown to late log phase. Protein expression was induced by addition of 0.5 mM isopropylthio- β -D-galactoside (IPTG). After another 3 h of growth at 37 or 30 °C, the bacteria

were harvested by centrifugation, and the cells were lysed by glacial acetic acid as described previously (16). Peptides were purified from the soluble fraction to homogeneity by reversed-phase high-performance liquid chromatography (HPLC), using a Vydac C-18 preparative column and a linear gradient of acetonitrile containing 0.1% trifluoroacetic acid. The C34 and C34/Q652L peptides were synthesized by solid-phase Fmoc methods and purified by HPLC as described above. In each peptide, the N-terminus is acetylated and the C-terminus is amidated. Peptide identity was confirmed by mass spectrometry. Peptide concentrations were determined by absorbance at 280 nm in the presence of 6 M guanidinium hydrochloride (GdmCl) (37).

Circular Dichroism (CD) Spectroscopy. CD spectra were acquired at 10 μ M peptide concentration in 50 mM sodium phosphate, pH 7.0, and 150 mM NaCl (PBS) with an Aviv 62 DS spectrometer as described previously (38). The wavelength dependence of molar ellipticity, $[\theta]$, was monitored at 0 or 20 °C as the average of five scans, using a five-second integration time at 1.0 nm wavelength increments. Spectra were baseline-corrected against the cuvette with buffer alone. Helix content was estimated from the CD signal by dividing the mean residue ellipticity at 222 nm by the value expected for 100% helix formation by helices of comparable size, $-33\,000\text{ deg cm}^2\text{ dmol}^{-1}$ (39). Thermal stability was determined by monitoring the change in CD signal at 222 nm as a function of temperature, and thermal melts were performed in two-degree intervals with a 2 min equilibration at the desired temperature and an integration time of 30 s. Reversibility was checked by repeated scans. The midpoint of the thermal unfolding transition (apparent T_m) was determined from the maximum of the first derivative, with respect to the reciprocal of the temperature, of the $[\theta]_{222}$ values (40). The error in estimation of T_m is ± 0.5 °C.

Sedimentation Equilibrium. Sedimentation equilibrium analysis was performed on a Beckman XL-A analytical ultracentrifuge as described previously (38). Peptide solutions were dialyzed overnight against PBS, loaded at initial concentrations of 10, 30, and 100 μ M, and analyzed at rotor speeds at 20 and 23 krpm at 20 °C. Data sets were fitted simultaneously to a single-species model with the program NONLIN (41). Protein partial specific volume and solvent density were calculated with constants from Laue et al. (42). Molecular weights were all within 10% of those calculated for an ideal trimer, with no systematic deviation of the residuals.

Crystallization, Data Collection, and Structure Determination. The Q652L variant was crystallized by the sitting drop vapor diffusion method at room temperature. To grow crystals, a 10 mg/mL HPLC-purified peptide stock was diluted 1:1 with a reservoir and allowed to equilibrate against the reservoir solution. Initial crystallization conditions were screened by using sparse matrix crystallization kits (Crystal Screen I and II, Hampton Research, Riverside, CA) and then optimized. Crystals of Q652L in space group *R*3 ($a = b = 52.38$ Å, $c = 60.48$ Å) were grown from 0.2 M ammonium sulfate, 0.1 M sodium acetate, pH 4.6, and 11% poly(ethylene glycol) 4000 and transferred to a cryoprotected solution containing 15% (vol/vol) glycerol in the corresponding mother liquor. Data were collected at 95 K using an R-axis IV image plate detector mounted on a Rigaku RU200 rotating anode X-ray generator at the X-ray Crystallography Facility

Table 1: Summary of Circular Dichroism and Sedimentation Equilibrium Data for N34(L6)C28 Mutants

| core model | $[-\theta]_{222}$ (deg cm ² dmol ⁻¹) | T_m (°C) | T_m^{GdmCl} (°C) ^a | molecular mass (kDa) |
|-------------|--|---------------|------------------------------------|-------------------------|
| N34(L6)C28 | 31 300 | 70 | 46 | 24.4 |
| Q652L | 31 900 | 72 | 48 | 24.9 |
| Q652L/1573S | 19 900 | 24 | | |

^a T_m^{GdmCl} denotes the melting temperature in 2 M GdmCl. All scans and melts were performed at 10 μ M peptide concentration.

Table 2: X-ray Data Collection and Refinement Statistics

| | |
|---------------------------------------|---------------------|
| unit cell dimensions | |
| a, b, c (Å) | 52.38, 52.38, 60.48 |
| α, β, γ (deg) | 90, 90, 120 |
| space group | R3 |
| data processing | |
| resolution (Å) | 50.0–2.0 |
| measured reflections | 23 565 |
| unique reflections | 4167 |
| completeness (%) | 99.7 |
| R_{merge} (%) | 4.0 |
| refinement | |
| resolution (Å) | 50.0–2.0 |
| reflections in working set | 3668 |
| reflections in test set | 499 |
| protein nonhydrogen atoms | 510 |
| water molecules | 77 |
| R_{free} (%) | 24.3 |
| R_{cryst} (%) | 20.6 |
| average B -factor (Å ²) | 38.5 |
| rms deviations from ideality | |
| bond lengths (Å) | 0.005 |
| bond angles (deg) | 0.768 |
| torsion angles (Å) | 14.8 |

^a $R_{merge} = \sum |I - \langle I \rangle| / \sum I$, where I is the intensity of an individual measurement and $\langle I \rangle$ is the average intensity from multiply recorded reflections. ^b $R_{cryst} = \sum |F_{obs} - F_{calc}| / \sum F_{obs}$, where F_{obs} and F_{calc} are observed and calculated structural factors, respectively. No σ -cutoff was applied. R_{free} is calculated for a set of reflections that were excluded from refinement.

at the Weill Medical College of Cornell University. Diffraction intensities were integrated by using DENZO and SCALEPACK software (43) and reduced to structural factors with the program TRUNCATE from the CCP4 program suite (44).

The structure of Q652L was determined by molecular replacement by using the program AMoRe (45). The 2.4 Å structure of wild-type N34(L6)C28 (PDB file name 1STZ) was used in a combined rotation-translation search (with 15.0–3.5 Å data) to yield a solution for Q652L (correlation coefficient = 66.1%; R -factor = 46.5%). The model was refined with simulated annealing and atomic displacement parameter (ADP) refinements using the program X-PLOR (46). Density interpretation and model building were done with the program O (47). Prior to refinement, 5% of the diffraction data were set aside for cross validation (free R -factor calculation). Crystallographic refinement of the structure was done with the program CNS 0.5 (48). The final refined model (R_{cryst} = 20.6%; R_{free} = 24.3%) includes 189–204 residues in the N34(L6)C28 trimer, and 77 water molecules (Table 2). Residues 546, 654, and 655 of gp41 and Gly 2 and Gly 3 of the linker were left out of the model because of the absence of interpretable electron density for these atoms. The Arg 579 side chain of gp41 and the Arg 4 side chain of the linker are disordered and were thus modeled as alanine.

Cell–Cell Fusion and Infectivity Assays. The inhibitory activities of N34(L6)C28 and variants thereof were determined by using a dye transfer fusion assay as described previously (49). HIV-1_{IIIB}-infected H9 cells were labeled with 2',7'-bis-(2-carboxyethyl)-5-(and-6)-carboxyfluorescein-acetoxymethyl ester (BCECF-AM). Fluorescently labeled HIV-1_{IIIB}-infected H9 cells (1×10^4) and MT-2 cells were cocultured at 1:10 ratio at varying peptide concentrations for 2 h at 37 °C in a 96-well microplate to obtain a dose–response curve. The fused cells were scored for dye transfer under fluorescence microscopy. In the infectivity assay as described previously (50), HIV-1_{IIIB} was inoculated with 1×10^4 of MT-2 cells at 0.0045 multiplicity of infection in RPMI 1640 containing 10% fetal bovine serum in the presence of varying peptide concentrations. After incubation at 37 °C for 1 and 24 h, half of the culture media was changed. HIV-1_{IIIB}-mediated cytopathic effect (CPE) was assessed by viability assay 6 days after infection.

RESULTS

Solution Properties of Q652L. Genetic studies indicate that substituting leucine for a conserved glutamine (Gln 652) within the C-terminal helix of gp41 results in higher than wild-type membrane fusion activity (25). This enhanced fusion activity is unlikely to be due to higher protein expression because the wild-type and mutant envelope glycoproteins have similar cell-surface expression (25). The crystal structure of the gp41 ectodomain core revealed that the Gln 652 side chain packs into a highly conserved hydrophobic groove of the coiled coil and that the carboxyl group of Gln 652 forms a tertiary hydrogen bond with the amide of Asn 554 (17–19). To understand how the Gln 652 to Leu mutation affects the structure and stability of the six-helix bundle, we substituted Gln 652 with leucine in the recombinant N34(L6)C28 model, a system that retains many structural features characteristic of the fusion-active gp41 core (19, 21). This mutant peptide was named Q652L.

The Q652L mutant peptide was produced by bacterial expression and purified to homogeneity by HPLC (see Materials and Methods). The circular dichroism spectrum of Q652L indicates that the folded mutant peptide contains greater than 95% α -helical structure (Figure 2a; Table 1). Sedimentation equilibrium measurements demonstrate that Q652L forms a clean trimer over a 10-fold range of peptide concentrations (10–100 μ M) (Figure 2c; Table 1). These results indicate that, like the wild-type peptide, the Q652L mutant forms a six-helix bundle structure in solution. Thus, the Gln 652 to Leu mutation does not affect the overall protein fold of the gp41 core.

The stability of the Q652L mutant peptide was assessed by monitoring the change in $[\theta]_{222}$ as a function of temperature. In neutral pH phosphate-buffered saline (PBS) at 10 μ M peptide concentration, the apparent melting temperature (T_m) of Q652L is 72 °C, relative to 70 °C for N34(L6)C28 (wild-type) (Table 1). Like the wild-type peptide, the thermally induced unfolding transition of the Q652L mutant is not reversible and the solution becomes turbid upon heat denaturation. In the presence of the denaturant guanidinium hydrochloride (GdmCl) at 1.5 M concentration, both the wild-type and mutant peptides undergo cooperative and reversible unfolding transitions, with T_m s of 54 and 56 °C,

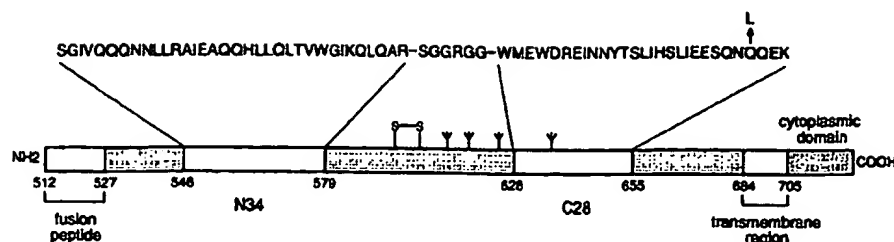


FIGURE 1: Schematic view of the HIV-1 gp41 envelope protein. The amino acid sequences of the N34 and C28 segments are shown. The N34(L6)C28 model of the gp41 ectodomain core consists of N34 and C28 plus a six-residue linker (21). A Gln 652 to Leu mutation that increases HIV-1 infectivity is indicated above the sequence. The disulfide bond and four potential N-glycosylation sites are depicted. The residues are numbered according to their position in gp160.

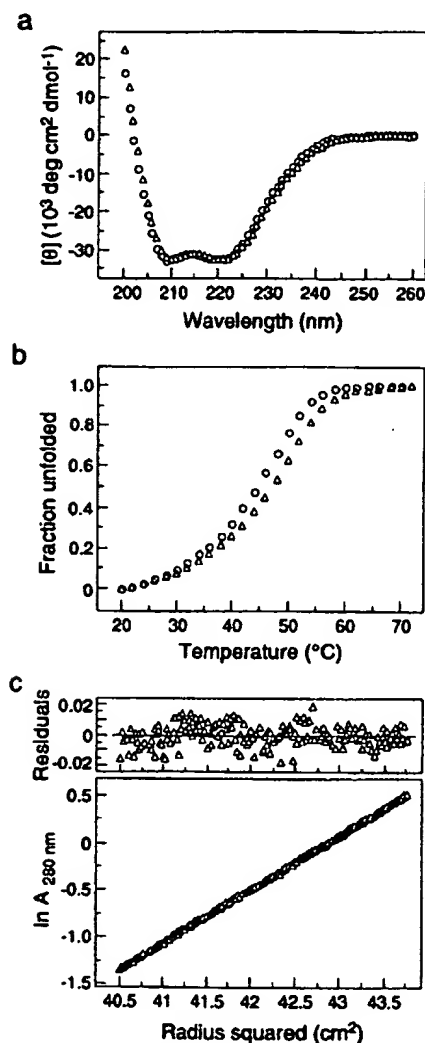


FIGURE 2: Folding of Q652L into a stable α -helical trimer. (a) Circular dichroism (CD) spectra of wild-type N34(L6)C28 (circles) and Q652L (triangles) at 0 °C in PBS (pH 7.0). (b) Thermal melts of N34(L6)C28 (circles) and Q652L (triangles) monitored by CD at 222 nm at 10 μ M peptide concentration in PBS (pH 7.0) in the presence of 2 M GdmCl. The decrease in the fraction of a folded molecule is shown as a function of temperature. (c) Analytical ultracentrifugation data (20 krpm) for Q652L collected at 20 °C in PBS (pH 7.0) at \sim 30 μ M peptide concentration. The natural logarithm of the absorbance at 280 nm is plotted against the square of the radial position. Deviations from the calculated values are plotted as residuals (top portion of panel c).

respectively (Figure 2b; Table 1). The change in thermal stability, ΔT_m , between the wild-type and mutant peptides

is 2 °C in the presence and absence of GdmCl. By using the Gibbs-Helmholtz equation, together with the thermodynamic parameters determined for the SIV N34(L6)C28 peptide (I. Jelesarov, H. R. Bosshard, H. J., and M. L., unpublished data), the corresponding free energy difference is -0.8 ± 0.3 kcal/mol for Q652L, relative to the wild-type peptide. Thus, the six-helix bundle structure of the gp41 core is stabilized by the Gln 652 to leucine substitution.

Crystal Structure of Q652L. To investigate the basis for the enhanced stability of the mutant molecule, the X-ray crystal structure of the Q652L peptide was determined to 2.0 Å resolution using the molecular replacement method. The final $2F_o - F_c$ electron density map is of good quality (see Figure 3) and reveals the positions of all the amino acid residues except four at the helix termini and in the linker region (Figure 4a). Omit maps were calculated to confirm the assignment of the side chains. The structure of Q652L was refined to a conventional R -factor of 20.6% with a free R -factor of 24.3% and root-mean-square deviations from ideal bond lengths and bond angles of 0.005 Å and 0.768°, respectively (Table 2). The quality of the refined model was checked by PROCHECK (57). All main-chain torsional angles fall within the helical region of a Ramachandran plot. Details of the data collection and refinement statistics are given in Table 2.

Like the wild-type peptide, the Q652L mutant folds into a six-helix bundle composed of three helical hairpins, each consisting of an N34 helix paired with an antiparallel C28 helix (Figure 4a). At the center of the bundle, the side chains at the heptad *a* and *d* positions of the N34 coiled coil displays typical "knobs-into-holes" packing interactions (52, 53). Three C28 helices wrap in the reverse direction against the outside of the N34 coiled-coil trimer and are tilted relative to the N34 superhelix by approximately 13° (Figure 4b). These C28 helices interact with the N34 helices through residues in three highly conserved hydrophobic grooves on the surface of the coiled-coil trimer (Figure 4c). In general, residues in the *a* and *d* positions of the C28 helix pack against residues at the *e* and *g* positions of the N34 helices. The root-mean-square (rms) deviation between all C_α atoms of the central N34 coiled coil in the wild-type and mutant molecules is 1.14 Å. The C28 helices in Q652L can also be superimposed upon the wild-type counterpart with a rms deviation of 1.52 Å. Thus, the overall architecture and helix packing of the Q652L mutant are the same as that of the wild-type molecule.

Although most of the helix packing interface within the gp41 core structure involves hydrophobic interactions, interhelical hydrogen bonds and salt bridges are uniformly

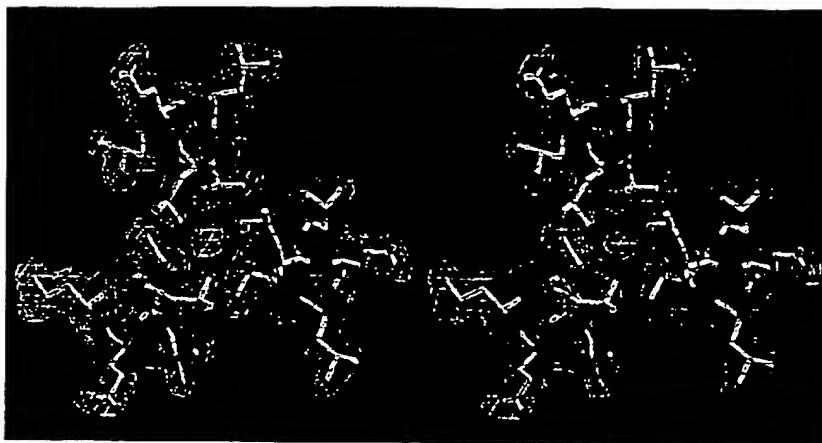


FIGURE 3: Stereoview of the final $2F_o - F_c$ electron-density map of the Q652L mutant with the refined model superimposed. Oxygen, nitrogen, and carbon atoms are colored red, blue, and white, respectively. A water molecule within the central N34 coiled coil of Q652L is shown as red balls. The map is contoured at 1.0 standard deviation above the average density. Figure was generated by using the program SETOR (66).

Table 3: Inhibitory Potency of gp41 Peptides

| peptide | IC ₅₀ (nM) | |
|-------------|-------------------------------|------------------------------|
| | cell-cell fusion ^a | HIV-infectivity ^b |
| N34(L6)C28 | 1508 ± 160 | 1487 ± 87 |
| Q652L | 180 ± 25 | 116 ± 12 |
| Q652L/I573S | 208 ± 17 | 308 ± 30 |
| I573S | 685 ± 9 | 730 ± 20 |
| C28 | 947 ± 67 | 2,192 ± 19 |
| C34 | 29.8 ± 1.7 | 17.2 ± 1.6 |
| C34/Q652L | 7.4 ± 0.7 | 6.8 ± 0.3 |

^a Inhibition of cell-cell fusion was measured in a syncytium formation. Indicated is the mean and standard deviation ($n = 4$).

^b Inhibition of HIV-1 infectivity was measured in a cell-culture infection assay. Indicated is the mean and standard deviation ($n = 3$).

distributed along the hydrophobic contacts. In the wild-type structure, a cluster of polar glutamines, including Gln 652, is involved in an intricate network of interhelical hydrogen bonds; a conserved salt bridge is formed between Asn 554 from N34 and Gln 652 from C28 in the wild-type structure (Figure 5). The substituted Leu 652 side chain forms a well-packed interface that lacks an interhelical hydrogen bond (Figure 5). The distances between the Leu 652 C β atom and the symmetry-related C β atoms of Gln 550 and Gln 551 from the N34 helix are 6.17 and 4.68 Å, respectively. In contrast, the corresponding distances for the wild-type Gln 652 side chain are 6.50 Å for Gln 550 and 5.73 Å for Gln 551. These favorable van der Waals interactions in the Q652L mutant appear to strengthen helix-helix packing and stabilize the six-helix bundle structure. In addition, the Gln 652 to Leu substitution also leads to local structural rearrangements; the side chains of Asn 553, Trp 628, and Gln 653 deviates substantially (Figure 4, panels b and c). These side chains may be poorly ordered due to chain termini effects or on the surface of the bundle. Overall, the nonpolar Leu 652 side chain makes good interhelical packing interactions and is well accommodated without significantly altering the six-helix bundle structure.

Antiviral Activity of Q652L. Synthetic peptides that make up the gp41 core are potent inhibitors of HIV infection and syncytium formation (49, 54, 55). We have previously shown that the N34(L6)C28 peptide itself inhibits HIV-1-mediated membrane fusion at micromolar concentrations (21). To

examine whether the fusion-enhanced Gln 652 to leucine mutation affects the antiviral activity of N34(L6)C28, we determined the relative potencies of the wild-type and mutant peptides to inhibit HIV-1 infection using both cell-cell fusion and infectivities assays. Figure 6 shows the inhibition of HIV-1_{IIIB}-infected H9 cell-mediated syncytium formation and cytopathic effect by different concentrations of peptides. Interestingly, the Q652L peptide is a highly effective inhibitor, with 50% inhibition concentrations (IC₅₀) of 0.18 μM for syncytium formation and 0.12 μM for viral entry, compared with IC₅₀ values of 1.51 and 1.49 μM for the wild-type N34(L6)C28 peptide (Figure 6; Table 3). Thus, the Gln 652 to leucine mutation increases the inhibitory activity of N34(L6)C28 about 10-fold.

To test if the six-helix bundle structure is required for the enhanced antiviral activity of the mutant peptide, we substituted Ile 573, an α heptad position, with Ser in Q652L to produce the double mutant Q652L/I573S. The Ile 573 to Ser mutation was chosen because it essentially disrupts the six-helix complex formation (21). Circular dichroism experiments confirm that this double mutant contains ~60% helical structure, with an apparent T_m of 24 °C (Figure 7a; Table 1). The Q652L/I573S peptide inhibits cell-cell fusion and infection with IC₅₀s of 0.21 and 0.31 μM, respectively, compared with IC₅₀ values of 0.69 and 0.73 μM for the single I573S mutant (Figure 7b; Table 3). In addition, the isolated C28 peptide blocks syncytium formation and viral entry with IC₅₀ values of 0.95 and 2.19 μM (Table 3). Taken together, these results strongly suggest that the enhanced antiviral activity of the Q652L peptide, relative to the wild-type molecule, results from the unfolded C28 segment bearing the Gln 652 to Leu mutation.

To further clarify this relationship, synthetic peptides with blocked termini, corresponding to the C34 sequence and its variant containing the Gln 652 to Leu mutation, were generated. The C34 peptide identified by the protein dissection method spans residues 628–661 and contains the entire region of the C28 peptide (16). C34 was chosen because this 34-residue peptide is soluble in aqueous solution and has high inhibitory activity (16, 26). Whereas wild-type C34 showed IC₅₀ values of 29.8 and 17.2 nM for inhibition of syncytium formation and HIV-1 entry, respectively, the

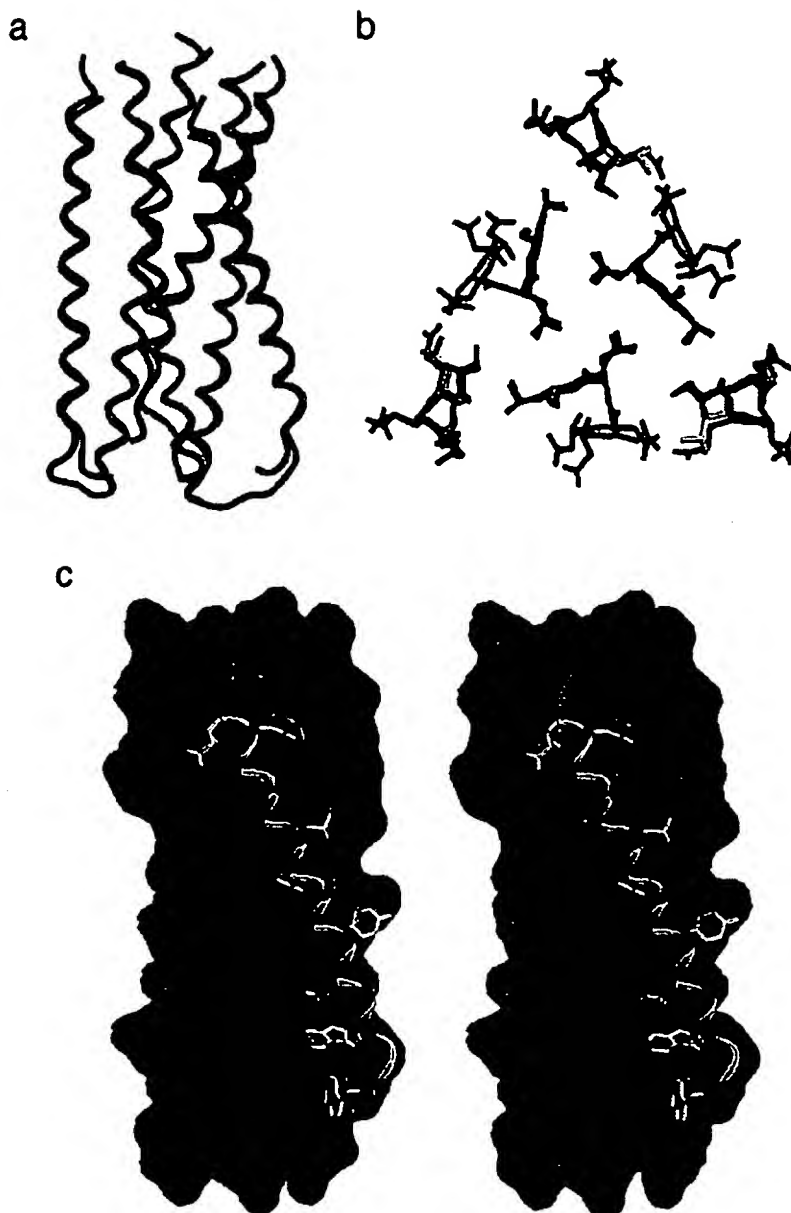


FIGURE 4: X-ray crystal structure of the Q652L mutant gp41 core. (a) Side view. The alpha-carbon traces of the N34 and C28 helices in the wild-type N34(L6)C28 (green) and mutant Q652L (red) structures were used for the superposition. The amino terminus of N34 and the carboxyl terminus of C28 are at the top of the figure. (b) A cross-section of the helix packing near the Gln 652 residue. Structures of wild-type N34(L6)C28 (green) and Q652L (red) are overlaid. The side chain of the mutated Gln 652 residue is colored yellow. (c) Stereoview of the interaction of the C28 helix with the conserved hydrophobic groove on the surface of the N34 coiled coil in the Q652L mutant gp41 core. The C28 helices of wild-type N34(L6)C28 (green) and Q652L (yellow), represented as ribbons, are shown against a surface representation of the central N34 coiled coil in wild-type N34(L6)C28. Figures were generated with the program SETOR (66).

IC₅₀ of C34/Q652L is 7.4 nM for syncytium formation and 6.8 nM for viral entry (Figure 8; Table 3). Thus, the Gln 652 to leucine mutation resulted in a ~2.5–4-fold increase in inhibitory activity.

DISCUSSION

The structure and mechanism of the HIV-1 envelope glycoprotein have been studied intensively. It has been shown that binding of gp120 to cell-surface receptors activates the membrane fusion activity of gp41, which then mediates membrane fusion (reviewed in ref 3). By analogy with the low-pH-induced structural changes in the influenza virus

hemagglutinin (HA) protein, this fusion activation process likely involves a substantial conformational change in the gp120–gp41 complex that switches from a labile native structure to an energetically more stable and fusogenic conformation (9, 10, 28, 56–59). The native gp120–gp41 structure is thought to be substantially different from that present in the fusogenic conformation, as has been established for the case of influenza HA (10, 60). Although little is known about the details of the receptor-activated conformational change in the HIV-1 envelope protein, the soluble ectodomain of gp41 is well characterized. The core structure of this ectodomain is a six-helix bundle (17–19, 61). Several

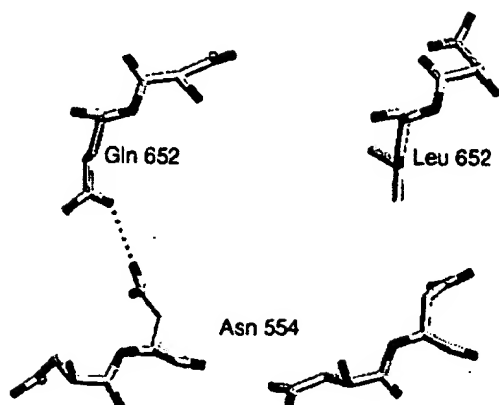


FIGURE 5: Side-chain carbonyl group of Gln 652 forms a tertiary hydrogen bond with the side-chain amide of Asn 554 in the wild-type structure (left panel). This hydrogen bond is not formed in the mutant Q652L structure (right panel). Oxygen, nitrogen, and carbon atoms are colored red, blue, and white, respectively. Hydrogen bond is shown in a purple dotted line. Figure was created with the program SETOR (66).

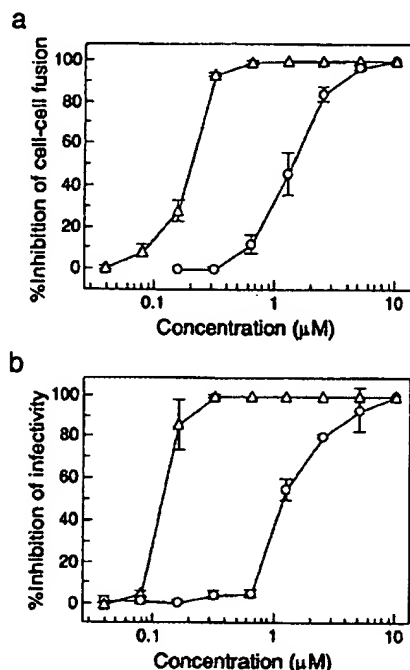


FIGURE 6: Inhibitory potency of the mutant Q652L peptide. (a) Inhibition of HIV-1_{IIIB}-infected H9 cell-induced cell-cell fusion by wild-type N34(L6)C28 (circles) and Q652L (triangles). Error bars indicate standard deviations from quadruplicate experiments. (b) Inhibition of HIV-1_{IIIB}-mediated cytopathic effect by wild-type N34(L6)C28 (circles) and Q652L (triangles). Error bars indicate standard deviations from triplicate experiments.

independent lines of evidence suggest that this structure corresponds to the fusogenic conformation of gp41 (15, 17–20, 24, 26, 27, 30, 32). Given that the fusion peptide region, which lies on the top of the N-terminal coiled coil, inserts into the cell membrane and the transmembrane helix at the C-terminus of the ectodomain is embedded in the viral membrane, the role of the six-helix bundle structure is proposed to facilitate the apposition of the viral and cellular membranes for fusion (6, 18, 23, 24). While few details are

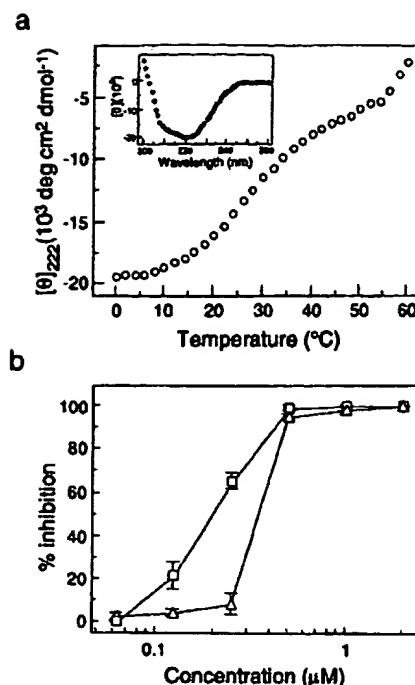


FIGURE 7: Ile 573 to Ser substitution in the Q652L mutant peptide disrupts the six-helix bundle formation and retains its potent inhibitory activity. (a) Thermal melt of Q652L/1573S monitored by CD at 222 nm. (Inset) CD spectrum at 0 °C at 10 μM peptide concentration in PBS (pH 7.0). (b) Inhibition of HIV-1_{IIIB}-infected H9 cell-induced cell-cell fusion (circles) and HIV-1_{IIIB}-mediated cytopathic effect (squares) by Q652L/1573S. Error bars indicate standard deviations from quadruplicate and triplicate experiments for cell-cell fusion and CPE, respectively.

known about how this membrane juxtaposition leads to complete membrane fusion, it is widely believed that the fusion involves the higher-order assembly of envelope protein trimers to form fusion pores (62).

Mutagenesis studies indicate that the packing interactions between the N- and C-terminal helices of gp41 are critical for HIV-1 membrane fusion. The conserved Gln 652 and Asn 656 residues on the buried face of the C-terminal helix are packed into a hydrophobic groove on the N-terminal coiled-coil surface within the gp41 core structure (17–19, 61). Whereas the Asn 656 to Leu mutation abolishes HIV-1 infectivity, the Gln 652 to Leu mutation enhances membrane fusion activity (25). Our X-ray crystallographic data indicate that this fusion-enhancement mutation sequence is well accommodated in the six-helix bundle structure by forming a packing surface with different sets of atoms. The crystal structures of the wild-type and mutant molecules can be superimposed, with a rms deviation of 1.40 Å. The free energy of helical interface stabilization associated with the Gln 652 to Leu mutation in the mutant structure is -0.8 kcal/mol, as determined by thermal unfolding studies. The enhanced membrane fusion activity by this mutation correlates with structural perturbations that strengthen helical-packing interactions within the six-helix bundle. Our results provide strong evidence that the helical-packing rearrangements triggered by receptor binding are likely to underlie the mechanism of the gp41 conformational change required for activation of HIV-1 membrane fusion. Interestingly, a similar mechanism is also proposed for many other viral membrane fusion proteins as well as the SNARE proteins

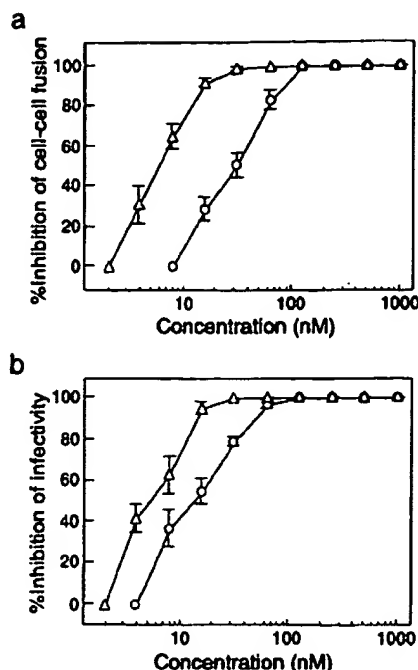


FIGURE 8: Inhibitory potency of the C34 and C34/Q652L peptides. (a) Inhibition of HIV-1-infected H9 cell-induced cell-cell fusion by C34 (circles) and C34/Q652L (triangles). Error bars indicate standard deviations from quadruplicate experiments. (b) Inhibition of HIV-1-mediated cytopathic effect by C34 (circles) and C34/Q652L (triangles). Error bars indicate standard deviations from triplicate experiments.

that mediate cellular membrane fusion processes (for a recent review, see ref 63).

Peptides corresponding to the N- and C-terminal heptad-repeat regions of gp41 are themselves effective inhibitors, at micromolar to nanomolar concentrations, of HIV-1 infection and syncytium formation (15, 49, 54, 55). Peptides from the C-terminal region of the ectodomain have the highest activity. It is striking that these peptides are effective against both laboratory-adapted strains and primary isolates of HIV-1 (49, 54, 55). The C peptide T-20 shows high efficacy in reducing viral loads in clinical trials (64). A large body of data is consistent with the proposal that these peptide inhibitors work through a dominant-negative mechanism by binding to viral gp41 and leading to its inactivation (15, 17–19, 24, 26–32). Kinetic studies demonstrate that the T-20 peptide does not act on the native gp120–gp41 complex but binds to gp41 only after interaction of the envelope complex with its cellular receptor (CD4) and coreceptors (e.g., CCR5 or CXCR4), suggesting strongly that the trimeric coiled coil of gp41 is available transiently during the conformational change to the fusogenic state (24). In addition, the conserved hydrophobic cavity on the gp41 coiled coil that provides a binding pocket for conserved C-peptide residues is proposed to be a good drug target for the development of small-molecule inhibitors of HIV-1 infection (26).

The N34(L6)C28 peptide model of the gp41 core is also capable of blocking HIV-1 membrane fusion at micromolar concentrations (21). Given that this inhibitory activity cannot be traced to the dissociated C28 region of the molecule, a mechanism different from that in the dominant-negative model proposed for the isolated N and C peptides is probably

involved in the inhibition of membrane fusion by the N34(L6)C28 complex (21). The inhibitory potency of N34(L6)C28 is dramatically increased by the fusion-defective Leu 568 to Ala and Trp 571 to Arg mutations that introduce structural perturbations in the coiled-coil cavity and thus weaken inter-helical packing interactions within the gp41 core (25, 27). This enhanced inhibitory activity is likely to result from the synergistic inhibition of the N34 and C28 peptides in the mutant molecules (27).

In this study, we demonstrate that the Gln 652 to Leu mutation also exerts striking effects on the inhibitory potency of N34(L6)C28; the Q652L mutant exhibits ~10-fold greater activity than the wild-type peptide. We also demonstrate that this enhanced inhibitory activity is likely to reflect that of the corresponding mutant C28 segment. Thus, the Gln 652 to Leu mutation can impart strong helical character and binding energy to the C28 peptide for binding to the trimeric coiled coil of gp41, as predicted by a dominant-negative mode of inhibition. Our data further show that the Gln 652 to Leu substitution can render the synthetic C34 peptide ~3-fold greater activity. Since this increase in inhibitory potency is less dramatic than that seen in the N34(L6)C28 variant, it is possible that more than one mechanism is responsible for the enhanced inhibitory activity of the Q652L peptide. Further studies are needed to ascertain the mechanism by which the six-helix bundle blocks viral fusion with host cells.

Overall, our results indicate that the conserved interhelical packing interactions within the six-helix bundle structure of gp41 play a role in HIV-1 entry and its inhibition. It remains to be determined whether these interactions underlie the mechanism for the proposed gp41 conformational change during fusion activation (6, 18, 23, 24; see, however, ref 61). Because the interhelical interactions have been shown to account for the broad antiviral activity of C peptides (55, 65), our results suggest that more potent analogues of C peptide inhibitors can be tested for the ability to form a stable complex with the N-terminal coiled coil of gp41.

REFERENCES

1. Turner, B. G., and Summers, M. F. (1998) *J. Mol. Biol.* 285, 1–32.
2. Luciw, P. A. (1996) in *Fields Virology* (Fields, B. N., Knipe, D. M., Howley, P. M., Chanock, R. M., Melnick, J. L., Monath, T. P., Roizman, B., and Straus, S. E., Eds.) pp 1881–1952, Lippincott-Raven Publishers, Philadelphia.
3. Moore, J. P., Jameson, B. A., Weiss, R. A., and Sattentau, Q. J. (1993) in *Viral Fusion Mechanisms* (Bentz, J., Ed.) pp 233–289, CRC Press, Boca Raton.
4. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) *Nature* 393, 648–659.
5. Rizzuto, C. D., Wyatt, R., Hernandez-Ramos, N., Sun, Y., Kwong, P. D., Hendrickson, W. A., and Sodroski, J. (1988) *Science* 280, 1949–1953.
6. Chan, D. C., and Kim, P. S. (1998) *Cell* 93, 681–684.
7. Stegmann, T., Delfino, J. M., Richards, F. M., and Helenius, A. (1991) *J. Biol. Chem.* 266, 18404–18410.
8. Tsurudome, M., Gluck, R., Graf, R., Falchetto, R., Schaller, U., and Brunner, J. (1992) *J. Biol. Chem.* 267, 20225–20232.
9. Carr, C. M., and Kim, P. S. (1993) *Cell* 73, 823–832.
10. Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994) *Nature* 371, 37–43.
11. Hernandez, L. D., Hoffman, L. R., Wolfsberg, T. G., and White, J. M. (1996) *Annu. Rev. Cell Dev. Biol.* 12, 627–661.
12. Chambers, P., Pringle, C. R., and Easton, A. J. (1990) *J. Gen. Virol.* 71, 3075–3080.

13. Delwart, E. J., Mosialos, G., and Gilmore, T. (1990) *AIDS Res. Hum. Retroviruses* 6, 703–706.
14. Gallaher, W. R., Ball, J. M., Garry, R. F., Griffin, M. C., and Montelaro, R. C. (1989) *AIDS Res. Hum. Retroviruses* 5, 431–440.
15. Lu, M., Blacklow, S. C., and Kim, P. S. (1995) *Nat. Struct. Biol.* 2, 1075–1082.
16. Lu, M., and Kim, P. S. (1997) *J. Biomol. Struct. Dyn.* 15, 465–471.
17. Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) *Cell* 89, 263–273.
18. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) *Nature* 387, 426–430.
19. Tan, K., Liu, J., Wang, J., Shen, S., and Lu, M. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 12303–12308.
20. Jiang, S., Lin, K., and Lu, M. (1998) *J. Virol.* 72, 10213–10217.
21. Lu, M., Ji, H., and Shen, S. (1999) *J. Virol.* 73, 4433–4438.
22. Dubay, J. W., Roberts, S. J., Brody, B., and Hunter, E. (1992) *J. Virol.* 66, 4748–4756.
23. Hughson, F. M. (1997) *Curr. Biol.* 7, R565–R569.
24. Furuta, R. A., Wild, C. T., Weng, Y., and Weiss, C. D. (1998) *Nat. Struct. Biol.* 5, 276–279.
25. Cao, J., Bergeron, L., Helseth, E., Thali, M., Repke, H., and Sodroski, J. (1993) *J. Virol.* 67, 2747–2755.
26. Chan, D. C., Chutkowski, C. T., and Kim, P. S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 15613–15617.
27. Ji, H., Shu, W., Burling, F. T., Jiang, S., and Lu, M. (1999) *J. Virol.* 73, 8578–8586.
28. Chen, C. H., Matthews, T. J., McDanal, C. B., Bolognesi, D. P., and Greenberg, M. L. (1995) *J. Virol.* 69, 3771–3777.
29. Wild, C. T., Greenwell, T., Shugars, D., Rimsky-Clarke, L., and Matthews, T. (1995) *AIDS Res. Hum. Retroviruses* 11, 323–325.
30. Judice, J. K., Tom, J. Y. K., Huang, W., Wrin, T., Vennari, J., Petropoulos, C. J., and McDowell, R. S. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 13426–13430.
31. Rimsky, L. T., Shugars, D. C., and Matthews, T. J. (1998) *J. Virol.* 72, 986–993.
32. Munoz-Barroso, I., Durell, S., Sakaguchi, K., Appella, E., and Blumenthal, R. (1998) *J. Cell Biol.* 140, 315–323.
33. Weissenhorn, W., Carfi, A., Lee, K.-H., Skehel, J. J., and Wiley, D. C. (1998) *Mol. Cell* 2, 605–616.
34. Malashkevich, V. N., Schneider, B. J., McNally, M. L., Milhollen, M. A., Pang, J. X., and Kim, P. S. (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96, 2662–2667.
35. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol.* 154, 367–382.
36. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Meth. Enzymol.* 185, 60–89.
37. Edelhoch, H. (1967) *Biochemistry* 6, 1948–1954.
38. Shu, W., Ji, H., and Lu, M. (1999) *Biochemistry* 38, 5378–5385.
39. Chen, Y. H., Yang, J. T., and Chau, K. H. (1974) *Biochemistry* 13, 3350–3359.
40. Cantor, C., and Schimmel, P. (1980) *Biophysical Chemistry*, Part III, pp 1131–1132, W. H. Freeman and Company, New York.
41. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) *Biophys. J.* 36, 575–588.
42. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.) pp 90–125, Royal Society of Chemistry, Cambridge.
43. Otwinowski, Z., and Minor, W. (1996) *Methods Enzymol.* 276, 307–326.
44. CCP4 (1994) *Acta Crystallogr., Sect. D* 50, 760–763.
45. Navaza, J. (1994) *Acta Crystallogr., Sect. A* 50, 157–163.
46. Brünger, A. T. (1992) *X-PLOR Version 3.1: A system for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
47. Jones, T. A., Zou, J.-Y., and Cowan, S. W. (1991) *Acta Crystallogr., Sect. A* 47, 110–119.
48. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr., Sect. D* 54, 905–921.
49. Jiang, S., Lin, K., Strick, N., and Neurath, A. R. (1993) *Nature* 365, 113.
50. Jiang, S., Lin, K., and Neurath, A. R. (1991) *J. Exp. Med.* 174, 1557–1563.
51. Laskowski, R. A., MacArthur, M. V., Moss, D. D., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291.
52. Crick, F. H. C. (1953) *Acta Crystallogr.* 6, 689–697.
53. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) *Science* 262, 1401–1407.
54. Wild, C. T., Oas, T., McDanal, C. B., Bolognesi, D., and Matthews, T. J. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 10537–10541.
55. Wild, C. T., Shugars, D. C., Greenwell, T. K., McDanal, C. B., and Matthews, T. J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 9770–9774.
56. Carr, C. M., Chaudhry, C., and Kim, P. S. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 14306–14313.
57. Stegmann, T., and Helenius, A. (1993) in *Viral Fusion Mechanisms* (Bentz, J., Ed.) pp 233–289, CRC Press, Boca Raton.
58. Wiley, D. C., and Skehel, J. J. (1987) *Annu. Rev. Biochem.* 56, 365–394.
59. Chen, J., Wharton, S. A., Weissenhorn, W., Calder, L. J., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92, 12205–12209.
60. Wilson, I. A., Skehel, J. J., and Wiley, D. C. (1981) *Nature* 289, 366–373.
61. Caffrey, M., Cai, M., Kaufman, J. Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. M., and Clore, G. M. (1998) *EMBO J.* 17, 4572–4584.
62. White, J. M. (1992) *Science* 258, 917–924.
63. Skehel, J. J., and Wiley, D. C. (1998) *Cell* 95, 871–874.
64. Kilby, J. M., Hopkins, S., Venetta, T. M., DiMassimo, B., Cloud, G. A., Lee, J. Y., Alldredge, Y., Hunter, E., Lambert, D., Bolognesi, D., Matthews, T., Johnson, M. R., Nowak, M. A., Shaw, G. M., and Saag, M. S. (1998) *Nat. Med.* 4, 1302–1307.
65. Malashkevich, V. N., Chan, D. C., Chutkowski, C. T., and Kim, P. S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 9134–9139.
66. Evans, S. V. (1993) *J. Mol. Graphics* 11, 134–138.

BI9921687

recognizing this gp41 core binds specifically to the surface of HIV-1 infected cells only after addition of soluble CD4 (8). This observation provides direct evidence that the trimer-of-hairpins structure represents the fusion-active state of gp41 (6). The structure of the HIV-1 envelope glycoprotein in its native conformation is unknown.

5 The viral envelope glycoprotein that mediates HIV-1 entry into target cells is an important target for humoral immunity during the natural course of infection (30). Current recombinant gp120 or gp41 protein vaccine candidates are unable to elicit antibodies capable of neutralizing primary HIV-1 isolates from infected persons at significant titers (34). These antibodies bind well to the individual gp120 and gp41 subunits but poorly to the membrane-associated, native envelope glycoprotein complex (Fig. 1) (35). In contrast, three infected
10 human-derived neutralizing antibodies have been shown to recognize the native gp120/gp41 complex efficiently (10). Thus, it appears that an immune response to virions, rather than to nonnative forms of the envelope glycoprotein (called viral debris), leads to the production of functional antibodies. Since gp120 is readily dissociated or shed from gp41 (37), the
15 challenge of inducing protective humoral immunity is to preserve the native envelope structure, or crucial components thereof, in vaccine preparations.

Because inhibition of gp41-mediated membrane fusion has potential to offer a general strategy for the treatment or prevention of HIV-1 infection (9), structural information on the conformational change of the protein is crucial to guiding efforts to target this process for
20 vaccine and antiviral drug development. In accordance with the present invention, it has been discovered that a peptide fragment derived from the COOH-terminal region of the gp41 ectodomain, forms a parallel three-stranded, α -helical coiled coil and serves as a trimerization domain in the native gp120/gp41 complex. A stable form of the trimeric coiled-coil domain elicits a strong antiviral antibody response against HIV-1 primary isolates from AIDS patients
25 and thus has utility as a safe and practical HIV-1 vaccine immunogen.

WMS:R P.9
OASIS-COAS

region of the ectodomain of enveloped viruses is easily discernible from the published sequences.

Similarly, conventional recombinant schemes may be used to obtain the isolated HIV gp41 trimer. In addition to the HIV Sequence Database and Medline, the complete nucleotide sequences of the HIV viruses had been previously reported by several investigators (Ratner et al., "Complete Nucleotide Sequence of the AIDS virus, HTLV-III," *Nature* 313:277-84 (1985); Muesing et al., "Nucleic Acid Structure and Expression of the Human AIDS/lymphadenopathy Retrovirus," *Nature* 313:450-458 (1985); Wain-Hobson et al., "Nucleotide Sequence of the AIDS Virus, LAV," *Cell* 40:9-17 (1985), which are hereby incorporated by reference). The RNA of the HIV-1 and HIV-2 viruses possess the following gene regions: so-called long terminal repeats at each end of the genome, *gag*, *pol*, *env*, and *nef*. The *gag* gene encodes the core proteins p24 and p17. The *pol* gene encodes the reverse transcriptase, RNase H, and integrase. The gene *nef* encodes a protein having a regulatory function. The *env* gene encodes the glycoproteins of the viral coat, namely gp41 and gp120. The *env* gene has been associated particularly with antigenicity and infectivity.

Given the knowledge of nucleic acid sequences encoding HIV, as well as other enveloped viruses, one of ordinary skill in the art can readily insert coding sequence for the gp41 glycoprotein or other viral envelope trimer comprising a C-terminal heptad repeat region of the ectodomain in a prefusogenic conformation, directly into a host cell or into an intermediate expression vector. Coding sequences corresponding to the HIV gp41 glycoprotein are readily available and may be obtained e.g., from the NIH AIDS Research and Reference Reagent Program Catalogue, Rockville, MD. Coding sequences for other enveloped viruses may be obtained from sources such as the American Type Culture Collection (ATCC), Manassas, Virginia.

Various truncated fragments of the nucleotide sequence encoding the gp41 glycoprotein may be used in accordance with the present invention. For example, coding sequence corresponding to amino acid residues 536-666, 640-658, or 624-669 (numbered according to their position in gp160, the HXB2 HIV-1 strain) may be used. Fragments of the gp41 glycoprotein may be obtained by truncating the 5', 3', or both 5' and 3' ends of the

DATA
SUPPORT
THIS
CONTRIBUTION

gp41 coding sequence by genetic manipulation. *In vitro* mutagenesis is helpful for introducing restriction sites. Similarly, truncated fragments of other enveloped virus C-terminal heptad repeat regions of the ectodomain may be used. There are various commercially available kits particularly suited for this application such as the T7-Gene *in vitro* Mutagenesis Kit (USB, Cleveland, OH) and the QuickChange Site Directed Mutagenesis Kit (Stratagene, San Diego, CA). Alternatively, PCR primers can be derived to allow direct amplification of a particular portion of a gp41 coding sequence or C-terminal heptad repeat region.

The carboxyl-terminal core region of the gp41 ectodomain (C-terminus trimeric coiled coil motif of an HIV gp41 ectodomain) may be stabilized by e.g., insertion of an isoleucine zipper. The C-terminal heptad-repeat regions of the ectodomain of other enveloped viruses may also be stabilized by insertion of an isoleucine zipper. Thus, for example, an isoleucine zipper described in P.B. Harbury, et al. (1994) *Nature* 371:80, may be inserted near the C-terminus of the coiled coil segment. Preferably, the isoleucine zipper is inserted in heptad register and before residue 684 in HIV (residues numbered according to their position in gp160, the HXB2 HIV-1 strain), which residue immediately precedes the transmembrane anchor. An isoleucine zipper may be inserted anywhere between residues 670 and 684 in HIV. Alternatively, an isoleucine zipper may be placed near the N-terminus of the C-terminal trimeric coiled coil of an HIV ectodomain or near the N-terminus of a C-terminal heptad-repeat region of the ectodomain of other enveloped viruses.

The carboxyl-terminal core region of the gp41 ectodomain may also be stabilized by cross linking of gp41 monomers. Similarly, the C-terminal heptad repeat regions of the ectodomain of other viral envelope proteins may be stabilized by cross linking. Such cross linking can be accomplished using well known methods such as those described in Houston, M.E., Jr. et al., (1995) *J. Peptide Science* 1:274-282.

The carboxyl-terminal core region of the gp41 ectodomain may also be stabilized by introducing one or point mutations in one or more gp41 monomers. Thus for example, Ser-649, Gln-652, Asn-656, and Glu-659 (numbered according to their position in gp160 in the HXB2 HIV-1 strain) may be mutated to the hydrophobic residue isoleucine. Methods of

same way that drug-resistance genes worked in bacteria, to allow rare transfected cells to grow up out of a much larger population that did not take up any DNA. The cells are transferred to selective growth medium, which permits growth only of cells that took up a functional tk gene (and the transferred DNA of interest). Various dominant selectable markers are now known in the art, including: aminoglycoside phosphotransferase (APH), using the drug G418 for selection which inhibits protein synthesis; the APH inactivates G418; dihydrofolate reductase (DHFR):Mtx-resistant variant, using the drug methotrexate (Mtx) for selection which inhibits DHFR; the variant DHFR is resistant to Mtx; hygromycin-B-phosphotransferase (HPH), using the drug hygromycin-B which inhibits protein synthesis; the HPH inactivates hygromycin B; thymidine kinase (TK), using the drug aminopterin which inhibits de novo purine and thymidylate synthesis; the TK synthesizes thymidylate; xanthine-guanine phosphoribosyltransferase (XGPRT), using the drug mycophenolic acid which inhibits de novo GMP synthesis; XGPRT synthesizes GMP from xanthine; and adenosine deaminase (ADA), using the drug 9-b-D-xylofuranosyl adenine (Xyl-A) which damages DNA; the ADA inactivates Xyl-A.

In addition to the recombinant DNA techniques described above for preparing a carboxyl-terminal core region of the gp41 ectodomain (C-terminus trimeric coiled coil motif of an HIV gp41 ectodomain), or C-terminal heptad repeat region of the ectodomain of another enveloped virus, such proteins may be synthesized using any number of well known techniques. For example, a carboxyl-terminal core region of the gp41 ectodomain may be prepared using the solid-phase synthetic technique initially described by Merrifield, in *J. Am. Chem. Soc.* 85:2149-2154 (1963). Other peptide synthesis techniques may be found in M. Bodanszky et al. *Peptide Synthesis*, John Wiley & Sons, 2d Ed. (1976) and other references readily available to skilled artisans. A summary of polypeptide synthesis techniques can be found in J. Stuart and J. D. Young. *Solid Phase Peptide Synthesis*, Pierce Chemical Company, Rockford, Ill., (1984). Peptides may also be synthesized by solution methods as described in *The Proteins*, Vol. II, 3d Ed., Neurath, H. et. al., Eds., p. 105-236, Academic Press, New York, N. Y. (1976). Appropriate protective groups for use in different peptide syntheses are described in the above-mentioned texts as well as in J. F. W. McOmie, *Protective Groups in Organic Chemistry*, Pelenu Press, New York, N.Y. (1973).

EXAMPLE 1

Construction of a chimera of gp41 and GCN4-pII. Formation of native envelope glycoprotein trimers on infectious virions is thought to be mediated by a sequence within the gp41 ectodomain other than the NH₂-terminal coiled-coil motif (Fig. 2A) (10). Trimerization therefore may therefore depend on a hydrophobic heptad repeat located in the COOH-terminal region of the protein. Earlier studies indicate the native structure of the gp120/gp41 complex is labile (11), and is transformed by receptor binding into an energetically more stable fusogenic conformation (5-7). Experiments were thus designed in order to test whether stabilizing this COOH-terminal coiled-coil motif could capture the gp41 molecule in the native metastable fold. Accordingly, the gp41 ectodomain, lacking the NH₂-terminal fusion peptide was engineered, by placing a soluble trimeric isoleucine zipper (GCN4-pII) in heptad register to the COOH terminus of the coiled-coil segment replacing the 17-residue region prior to the transmembrane anchor (Fig. 2B)

~~The genes of gp41 (residues 536 to 666, numbered according to their position in gp160 the HXB2 HIV-1 strain) and GCN4-pII (14) were subcloned into the expression vector pTMHa [J. P. Staley and P. S. Kim, *Protein Science* 3, 1822 (1994)] to produce plasmid pRgp41pII (Fig. 2B). Met⁶²⁹ in the pRgp41pII construct was mutated to leucine by single-stranded mutagenesis [T. A. Kunkel, J. D. Roberts, R. A. Zakour, *Methods Enzymol.* 154, 367 (1987)]. An expression vector (pC45pII) encoding the C45-pII peptide was derived from pRgp41pII by PCR amplification followed by subcloning into pTMHa. The *KpnI* site in pC45pII was replaced by the coding sequence for the residues Ala-Ser. Plasmid pC45 was derived from pC45pII. Plasmid pC52 was derived from pC45pII by the insertion of the appropriate DNA sequences encoding the residues Leu-Trp-Asn-Trp-Phe-Asn-Ile in place of GCN4-pII. Plasmid pC29p was derived from pC45pII. Single-letter abbreviations for amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.~~

In addition, Ile⁵⁷³ and Leu⁵⁷⁶ were mutated in this chimera molecule to polar serine residues in order to destabilize the NH₂-terminal coiled coil because it forms the centre of fusion-active gp41 (6, 12). The resulting chimera protein was designated rgp41-pII.

The rgp41-pII protein was produced by bacterial expression, purified by reverse-phase high-performance liquid chromatography, and refolded *in vitro*. Proteins were expressed in the *E. coli* strain BL21(DE3)/pLysS (Novagen). The chimeric proteins and peptides were purified essentially as described for the N36-p1 peptide [W. Shu, H. Ji, M. Lu, *Biochemistry* 38, 5378 (1999)]. Cys-containing proteins were subjected to air oxidation in 6 M guanidine hydrochloride (GdmCl), and 100 mM tris-HCl (pH 8.6) to permit disulfide bond formation prior to cyanogen bromide treatment. Purified rgp41-pII solubilized in 6 M GdmCl and 50 mM tris-HCl (pH 8.0) was refolded by dilution into 50 mM tris-HCl (pH 8.0). Identity of the peptides was confirmed by mass spectrometry (Perceptive Biosystems Voyager Elite), and all molecular masses were found to be within 2 daltons of the expected mass.

The rgp41-pII protein forms an insoluble aggregate under physiological conditions. Protease treatment with trypsin resulted in a soluble peptide fragment that extends from Ser⁶¹⁸ to the COOH-terminal end of the recgp41-pII construct, including the gp41 segment (residues 618 to 666) on the NH₂ terminus and the entire GCN4 coiled coil on the COOH terminus (Fig. 2C). To facilitate further studies, a molecule denoted C45-pII, was constructed in which the C45 segment (residues 624 to 668) of gp41 is fused in frame to GCN4-pII (Fig. 3A). The circular dichroism (CD) spectrum of C45-pII demonstrates that the peptide is fully helical in solution (Fig. 3B).

EXAMPLE 2

Alpha-helical structure of the chimera. In an attempt to define a smaller subdomain for high-resolution crystallographic studies, C45-pII was subjected to proteolysis by proteinase K. Trypsin (Worthington) or proteinase K (Boehringer Mannheim) digestions were performed with 1 mg/ml protein and 0.01 mg/ml at room temperature in 50 mM tris-HCl (pH 8.0) and were quenched with 4 mM phenylmethylsulphonyl fluoride (Sigma). The proteolytic fragments were separated and purified by reverse-phase high-performance liquid

the C45-pII peptide sediments as a trimeric species (Fig. 3C) (17). Thus, the chimera of the gp41 heptad repeat and isoleucine zipper can form a protease-resistant, three-helix structure.

EXAMPLE 3

Folding of the C52 coiled-coil domain. In contrast to the C45-pII chimera, the isolated C45 peptide (residues 624 to 669 of gp41) displays little secondary structure, as judged by CD spectroscopy (15). To investigate the possibility that the heptad-repeat motif might be stabilized in the context of the longer coiled coil, a 52-residue peptide, designated C52, that contains the entire C45 region plus 7 residues on the COOH terminus (Fig. 4A), was studied. The C52 peptide is ~70% helical and exhibits a reversible thermal unfolding transition at 100 μ M peptide concentration in neutral pH phosphate-buffered saline (PBS) at 0°C (Fig. 4B, C). Because the mutant envelope glycoprotein exhibits substantial cell-surface expression, gp160 precursor processing, CD4 binding, gp120/gp41 association, and soluble CD4-induced shedding (17), the fusion-defective Asn⁶⁵⁶ to Leu mutation is unlikely to alter structural features sufficiently to disrupt the native gp120/gp41 complex.

Sedimentation equilibrium measurements indicate that the C52 peptide is a trimer between 30 and 100 μ M (Fig. 4D). The molecular weight of the C52 peptide systematically decreases at concentrations below 25 μ M (presumably because of dissociation in the low micromolar concentration range). Thus, the C52 peptide forms a labile three-stranded coiled coil in isolation. The invariant Trp⁶⁷⁰ and Phe⁶⁷³ residues in positions *a* and *d* of the C52 peptide may contribute to the folding free energy of the coiled-coil trimer.

EXAMPLE 4

Model for the native conformation of gp41. The finding that the C52 peptide can fold into an α -helical trimer leads to a proposal of a model for the native conformation of gp41 (see Fig. 5). The COOH-terminal regions of the three gp41 chains form a parallel trimeric coiled coil that is ~9 residues from the viral membrane end of the molecule. The NH₂-terminal heptad-repeat regions make contacts with the gp120 subunits, and are thus held in a non-coiled-coil conformation (Fig. 5A). The gp120 chains may also contribute to the native structure but primarily form three membrane distal globular domains containing the

Halvoson, *Biophys. J.* **36**, 575 (1981)]. Solvent density and protein partial specific volume were calculated according to solvent and protein composition, respectively [T. M. Laue, B. D. Shah, T. M. Ridgeway, S. L. Pelletier, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, S. E. Harding, A. J. Rowe, J. C. Horton, Eds. (Royal Society of Chemistry, Cambridge, 1992), pp. 90-125]. The data fit well to a model for an ideal trimer, with no systematic deviation of the residuals. The observed molecular weight is 32,100 daltons, compared to an expected value of 31,400 daltons for a trimer.

Sedimentation equilibrium measurements indicate that the C52-pII peptide is a discrete trimer and exhibits no systematic deviation from trimer molecular weight between 5 and 50 μM (Fig. 7E). Thus, C52-pII forms a well-structured and extremely stable, three-stranded coiled coil.

Example 8

Eliciting neutralizing antibodies. To explore the immunogenicity of the COOH-terminal coiled-coil region of gp41, the C52-pII, C52, and C45-pII model peptides were used to immunize rabbits. To generate antisera, two rabbits each were immunized with the peptide C52-pII, C52 or C45-pII. The primary immunization was performed with a dose of 500 μg of peptide immunogen emulsified (1:1 mixture) in complete Freund's adjuvant. The animals were subsequently challenged with 250, 250, 125, and 125 μg of the immunogen with incomplete Freund's adjuvant for each of four booster immunizations at 3-week intervals. Sera collected 2 weeks after the third booster inoculation were used for the studies described here. Rabbit polyclonal antibodies (total IgG) were purified from serum on a protein A-Sepharose column (Amersham Pharmacia Biotech). Total IgG concentrations were determined by absorbance at 280 nm (1 optical density = 0.8 mg/ml). The rabbit antibodies were generated by Covance Research Products, Inc. (Denver, PA).

The C52 and C45-pII molecules were chosen as controls because the former is predominantly unfolded under physiological conditions and the latter lacks seven highly conserved hydrophobic amino acids at the COOH-terminus of the gp41 coiled-coil segment, relative to C52-pII (38).